Some statistical estimation problems in ruin theory

Susan M. Pitts

University of Cambridge, UK

email: `s.pitts@statslab.cam.ac.uk`

Abstract

Much research in ruin theory in insurance mathematics focuses on the behaviour of various quantities of interest, such as the probability of ruin or the ruin-time moments, for a particular risk model in insurance. In practice, precise knowledge of the risk model is available only via observed data. In this presentation, the problem of statistical estimation of the quantities of interest, given data on claim arrivals and claim sizes, is considered. Nonparametric estimators are proposed and their statistical properties are studied. In many cases, the quantities of interest are functions, and this leads to the use of techniques for estimation in function spaces. The bootstrap is used to obtain simultaneous confidence bands for the unknown quantities.

Key Words: Risk models, nonparametric estimation, bootstrap.

1   Introduction

In this paper, we consider statistical estimation problems for risk models in insurance mathematics. In practical applications of stochastic modelling in general, the parameters of the assumed underlying model are unknown and must be estimated from the available data, and this gives rise to statistical error in the estimation of input quantites, such as the claim-arrival rate or the claim-size distribution. One main purpose in using stochastic models for real-life situations is to use the model to learn about output quantities of interest, for example, the expected total amount of claims in a fixed period. In practice these quantites are often calculated as being those belonging to a model with input given by the estimated input quantities, which means that statistical error in the estimation of the input quantities is transferred into statistical error in the calculated output quantities. Investigation of the likely size of such error plays a role in the assessment and control of risk.

Here we provide an overview of one approach to such statistical estimation problems in risk and ruin theory which has been used for various quantities of interest in risk and ruin theory. In Section 2 we give a description of the classical risk model and define some ruin quantities that have been much studied in the literature. We consider statistical estimation in Section 3, with parametric examples to explain and motivate the approach, and then move on to the general set up for the nonparametric case in Section 4.
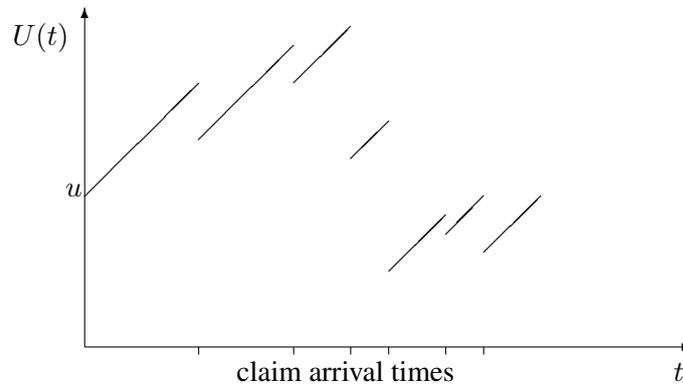
Figure 1: The surplus process $\{U(t) : t \geq 0\}$

2 Some ruin quantities in the classical risk model

Actuarial risk models may be used to model the surplus of a portfolio of insurance risks. One of the simplest models is the classical risk model, where claims arrive in a Poisson process rate λ, claim sizes are independent identically distributed (iid) positive random variables $X_1, X_2, \ldots$ with finite mean μ, independent of the claim arrivals process, and where premiums are received continuously in time at constant rate c (> 0). Initially, the surplus is $u \geq 0$, and, at time $t \geq 0$, the surplus is

$$U(t) = u + ct - S(t) = u + ct - \sum_{i=1}^{N(t)} X_i,$$

where $N(t)$ is the number of claims received in $(0, t]$ (so $N(t)$ has a Poisson distribution with mean λt) and $S(t) = \sum_{i=1}^{N(t)} X_i$ is the total amount of claims received by time t (so $S(t)$ is a compound Poisson random variable). We assume that $c > \lambda\mu$, ie that we have positive safety loading, and we write $c = (1+\theta)\lambda\mu$ where $\theta > 0$ is the premium loading factor. A typical sample path of the stochastic process $\{U(t) : t \geq 0\}$ is shown in Figure 1. The surplus increases at constant rate c between claim arrivals, and, when a claim arrives, the surplus jumps down by a distance equal to the amount of the arriving claim. One possible generalisation of the classical risk model is the Sparre Andersen model, where claims arrive in a renewal process, so that the inter-claim arrival times are iid random variables, but not necessarily exponentially distributed.

If $U(t) < 0$ for some $t > 0$ then ruin is said to occur. There are various quanitities connected with ruin that are of interest in the literature. The most obvious is $\psi(u) = \mathbb{P}(U(t) < 0 \text{ for some } t > 0)$, and with our assumption of positive safety loading we have $\psi(u) < 1$ (Corollary 1.4 of Chapter IV in Asmussen and Albrecher (2010)). There is also the finite-time ruin probability defined as $\psi(u, t_0) = \mathbb{P}(U(t) < 0 \text{ for some } t \in (0, t_0])$. Ruin probabilities have been much studied, giving rise to many theoretical results for them. One of these is the Cramér–Lundberg inequality (see, for example, §6.5 of

Dickson (2005)), which says that, under conditions on the moment generating function $M_X(r)$ of the claim-size distribution, we have $\psi(u) \leq e^{-Ru}$ for all $u \geq 0$, where R is the unique positive solution of $M_X(r) - 1 = (1+\theta)\mu r$. The quantity R is called the adjustment coefficient, and is another ruin quantity that has received much attention. The time to ruin is the defective random variable $\tau(u) = \inf\{t > 0 : U(t) < 0\}$, and is another key ruin quantity. We note that the probability of ruin $\psi(u)$ is $\mathbb{P}(\tau(u) < \infty)$, and the finite-time ruin probability is $\psi(u, t_0) = \mathbb{P}(\tau(u) \leq t_0)$. The moments of $\tau(u)$ are defined (under conditions on the moments of the claim-size distribution) as $\psi_k(u) = \mathbb{E}(\tau(u)^k 1(\tau(u) < \infty))$ for $k = 1, 2, \ldots$, where $1(A)$ is the indicator function of an event A, and the corresponding conditional moments are $\psi_k^C(u) = \mathbb{E}(\tau(u)^k | \tau(u) < \infty) = \psi_k(u)/\psi(u)$.

As is often the case in applied probability, these ruin quantities have easy closed form expressions only for certain choices of distribution for the claim sizes. For example, for claims that are exponentially distributed with mean μ, we have

$$R = \frac{1}{\mu} - \frac{\lambda}{c} = \frac{\theta}{(1+\theta)\mu},$$

$$\psi(u) = \frac{1}{1+\theta} e^{-Ru}$$

and

$$\psi_1^C(u) = \frac{1 + \theta + (u/\mu)}{(1+\theta)\theta\lambda}.$$

3   Statistical estimation in risk models

As discussed in the introduction, in real-life problems, the true values of the Poisson arrival rate and the parameters of the claim-size distribution are unknown and must be estimated from data. For example, for a classical risk model with exponentially distributed claim sizes with mean μ, suppose that λ is known (for example, from previous experience) but that μ is to be estimated from data $X_1, \ldots, X_n$, an iid sample of claim sizes. Suppose that we are interested in estimation of the adjustment coefficient R.

We use the above example to illustrate a possible approach, which is to use the data to estimate μ by, for example, the maximum likelihood estimator $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=0}^{n} X_i$, and then to estimate the adjustment coefficient R by $\hat{R} = \frac{1}{\hat{\mu}} - \frac{\lambda}{c}$. (This is a simplified description of the estimator; in fact we need a slightly more complicated definition which takes account of whether or not the data lead to $c > \lambda\hat{\mu}$.) This gives a point estimate $\hat{R}$ for the unknown R. It is important to go further, with the aim of quantifying the likely size of the statistical error in the point estimate. In this example, $\hat{R}$ is a relatively simple function of $\hat{\mu}$ so that it is straightforward to translate what is known about the statistical variation of the maximum likelihood estimator $\hat{\mu}$ into results for the statistical variation of $\hat{R}$, using the relevant function that maps $\hat{\mu}$ onto $\hat{R}$.

This example typifies the so-called "plug-in" approach where interest lies in a function $g(\mu)$ and data $X_1, \ldots, X_n$ are available that allow us to estimate μ. We then estimate $g(\mu)$ by the plug-in estimator $g(\hat{\mu})$, and combine known statistical properties of $\hat{\mu}$ as an estimator of μ with analytic properties of the map g in order to obtain statistical properties of the estimator $g(\hat{\mu})$. In the adjustment coefficient example, we can use the fact that $n\hat{\mu}$ has a gamma distribution to give exact confidence intervals for

the true R, and we can also make some headway in evaluating $\text{var}(\hat{R})$. However, for an approach that is applicable when such exact calculations are not possible, we may use the well-known delta method to translate the known asymptotic normality of the maximum likelihood estimator $\hat{\mu}$ into asymptotic normality of $\hat{R}$, provided g satisfies a differentiability condition. The resulting asymptotic normal distribution has variance of the form $\sigma^2(\mu)$, and we may estimate this by $\sigma^2(\hat{\mu})$, and use then use it for inference about R.

If λ is also unknown and is estimated, for example, by the maximum likelihood estimator $\hat{\lambda} = \bar{T}^{-1}$ based on an iid sample of inter-claim arrival times $T_1, \ldots, T_n$, then the above generalises to $R = g(\lambda, \mu)$ for some function $g : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$, and we may use the finite-dimensional delta method to derive an asymptotic normality result for $\hat{R}$ with asymptotic variance $\sigma^2(\lambda, \mu)$. Estimating this by $\sigma^2(\hat{\lambda}, \hat{u})$ then gives an approximate asymptotic measure of the variation in $\hat{R}$.

The two examples above illustrate the use of the finite-dimensional delta method in a parametric plug-in estimation problem. This may be further generalised to a nonparametric estimation problem, and we discuss this in the next section.

4   The nonparametric set up

In the nonparametric set up, we do not make distributional assumptions about the claim-size distribution, but instead assume that claim sizes are iid with unknown distribution function F. For example, if we are interested in the adjustment coefficient, and if for convenience we assume λ is known, then we regard the classical risk model as a map Φ that takes the claim-size distribution function F onto R, ie $R = \Phi(F)$, so Φ is a map with an appropriate function space for its domain (this example is considered in Hipp (1996), see also Csőrgö and Teugels (1990)). Generalising one step further, suppose that we are interested in the function $\psi_k(\cdot)$ (instead of R). Then we can think of the stochastic model as a map Ψ_k so that $\psi_k = \Psi_k(F)$ and Ψ_k is a map from one function space to another.

In general we have a stochastic model with data $X_1, \ldots, X_n$ available on an input distribution F (for example, the claim-size distribution), and a functional Φ that maps the input quantity onto an output quantity H of interest for which we want to have an estimator, ie

$$H = \Phi(F). \tag{4.1}$$

We use the data to construct an estimator $\hat{F}_n$ of F, and in our nonparametric examples, we use the empirical distribution function $\hat{F}_n$ where

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1(X_i \le x).$$

A plug-in estimator of H is

$$\hat{H}_n = \Phi(\hat{F}_n).$$

In this general set up, Φ is a map between Banach spaces. Let $(B_i, \|\cdot\|_i)$ be a Banach space with norm $\|\cdot\|_i$, $i = 1, 2$ and suppose that $\Phi : B_1 \to B_2$. For example, for (4.1), B_1 is an appropriate function space. The aim now is to establish asymptotic properties of $\hat{H}_n$ as an estimator of H, and for this we

hope to apply the infinite-dimensional delta method (see for example Gill (1989) and van der Vaart (1998)), which, in broad terms, states that, provided

(i) $\sqrt{n}(\hat{F}_n - F)$ converges in distribution to a Gaussian process E in $(B_1, \|\cdot\|_1)$ as $n \to \infty$, ie there is an empirical central limit theorem in $(B_1, \|\cdot\|_1)$), and

(ii) the map Φ satisfies an appropriate differentiability property,

then we have

$$\sqrt{n}\Big(\hat{H}_n - H\Big)\Big(= \sqrt{n}\Big(\Phi(\hat{F}_n) - \Phi(F)\Big)\Big) \text{ converges in distribution to } Z \text{ in } (B_2, \|\cdot\|_2),$$

where Z is a zero-mean Gaussian process in $(B_2, \|\cdot\|_2)$ whose sample paths are obtained by applying the derivative of Φ to the sample paths of the limit Gaussian process E in the empirical central limit theorem. The concept of convergence in distribution in B_i is as in Pollard (1984).

In order to apply this technology to estimation problems in risk and ruin theory, and to stochastic models more widely, it is necessary to choose the spaces B_1 and B_2 in such a way that, on the one hand, there is an empirical central limit theorem in B_1 and, on the other hand, that the map Φ satisfies the relevant differentiability property as a map from B_1 to B_2. There is a tension between these two requirements. Weighted spaces of right-continuous functions with left-hand limits provide a useful and flexible class of function spaces in statistical estimation problems such as those considered here. The relevant functional differentiability property for Φ is a version of Hadamaard differentiability (see, for example, van der Vaart (2000)). In general, the proof of the differentiability of a particular Φ throws up difficulties and challenges which must be solved separately on a case-by-case basis.

The covariance structure of the limiting output Gaussian process often turns out to depend in a complicated way on the input quantities, making it difficult to evaluate the variation in $\hat{H}_n$. However, bootstrap methods may be used to obtain a bootstrap confidence region in B_2 for the unknown quantity H of interest. The differentiability of the functional Φ typically leads to a formal justification that such confidence regions will have asymptotically the correct coverage rates.

The above approach has been applied to various quantities in risk theory, including compound distribution functions and the probability of ruin in the classical risk model (Pitts (1994)), the adjustment coefficient in the Sparre Andersen model (Pitts et al. (1996)), the survival probability $1 - \psi(\cdot)$ for the classical risk model where λ is also estimated (a semiparametric approach) (Politis (2003)), the survival probability $1 - \psi(0, \cdot)$ with zero initial capital in the classical risk model (Qin and Pitts (2012)), the ruin-time moments (Qin and Pitts (2013)).

Statistical estimation in risk and ruin theory has been considered by many authors, see, for example, Pitts (2004) and the references therein. These include other approaches to estimation in risk and ruin models, and there may also be other set ups, for example, we may have direct observations on different quantities, and these may mean that different statistical tools will be more relevant. Or we may have a different observation scheme, for example, we may observe the claim arrivals and sizes up to a fixed time T (see Grandell (1991)). Possible developments include the application of the methods discussed here to other risk models and other quantities of interest.

References

Asmussen, S., and Albrecher, H. (2010) *Ruin Probabilities*, 2nd edition, World Scientific Publishing, London.

Csőrgö, S., and Teugels, J.L. "Empirical Laplace transform and approximation of compound distributions," *Journal of Applied Probability* **27** 88–101.

Dickson, D.C.M. (2005) *Insurance Risk and Ruin*, Cambridge University Press, Cambridge.

Gill, R.D. (1989) "Non- and semi-parametric maximum likelihood estimates and the von Mises method (Part I)," *Scandinavian Journal of Statistics* **16** 97–128.

Grandell, J. (1991) *Aspects of risk theory*, Springer, New York.

Hipp, C. (1996) "The delta-method for actuarial statistics," *Scandinavian Actuarial Journal* 79–94.

Pitts, S.M. (1994) "Nonparametaric estimation of compound distributions with applications in insurance," *Annals of the Institute of Statistical Mathematics* **46** 537–555.

Pitts, S.M. (2004) "Estimation," in *Encyclopedia of Actuarial Science,* Vol. 2. eds Teugels, J.L., and Sundt, B. Wiley, 621–631.

Pitts, S.M., Grübel, R., and Embrechts, P. (1996) "Confidence bounds for the adjustment coefficient," *Advances in Applied Probability* **28** 802–827.

Politis, K. (2003) "Semiparametric estimation for non-ruin probabilities," *Scandinavian Actuarial Journal* 75–96.

Pollard, D. (1984) *Convergence of Stochastic Processes*, Springer–Verlag, New York.

Qin, L., and Pitts, S.M. (2012) "Nonparametric estimation of the finite-time survival probability with zero initial capital in the classical risk model," *Methodology and Computing in Applied Probability* **14** 919–936.

Qin, L., and Pitts, S.M. (2013) "Nonparametric estimation of the ruin time moments in the classical risk model," Submitted for publication.

van der Vaart, A.W. (2000) *Asymptotic Statistics*, Cambridge University Press, Cambridge.